Habitat International
Volume 50, December 2015, Pages 354–365
http://www.sciencedirect.com/science/article/pii/S0197397515001940
doi:10.1016/j.habitatint.2015.09.005

A multi-scale modeling approach for simulating urbanization in a metropolitan region

Saad Saleem Bhatti, Nitin Kumar Tripathi, Vilas Nitivattananon, Irfan Ahmad Rana, Chitrini Mozumder

Abstract

Metropolitan regions worldwide are experiencing rapid urban growth and the planners often employ prediction models to forecast the future expansion for improving the land management policies and practices. These regions are a mix of urban, peri-urban and rural areas where each sector has its unique expansion properties. This study examines the differences in urban and peri-urban growth characteristics, and their impact at different stages of prediction modeling, in city district Lahore, Pakistan. The analysis of multi-temporal land use/land cover maps revealed that the associations between major land transitions and the factors governing land changes were unique at city district, urban and peri-urban scales. A multilayer perceptron neural network was employed for modeling urbanization, and it was found that the sub-models developed for urban and peri-urban subsets returned better accuracies than those produced at the city district scale. The prediction maps of 2021 and 2035 were also produced through this approach.

Keywords: driving factors; land use/land cover change; multiple scenarios; neural network; peri-urban; urban growth modeling

1. Introduction

Urban growth is a complex process driven by a variety of spatio-temporal characteristics and a mixture of diverse components (Deng, Wang, Hong, & Qi, 2009). It takes place on a regional level and is usually hard to interpret and quantify (Jaeger, Bertiller, Schwick, & Kienast, 2010). Generally resulting in an increase in urban and decrease in rural areas, the land use/land cover (LULC) changes are governed by a myriad of choices like suitability of location, policies and individual preferences (Irwin & Bockstael, 2004). A metropolitan region includes urban, peri-urban (also referred as suburban) and rural areas; McGee (1995) mentioned that its development and growth must be dealt with as region- rather than city-based. A peri-urban area is generally defined as the transition zone between urban and rural areas possessing some characteristics of the both (Shi, Sun, Zhu, Li, & Mei, 2012). However, urban and peri-urban areas have their own trajectories and patterns of urbanization (Zanganeh Shahraki et al., 2011). Moreover, the factors governing land management and growth in peri-urban areas are somewhat different compared to the urban areas, and thus cannot be simultaneously used to understand the dynamics of the both.

Unplanned urban sprawl in metropolitan regions is a serious concern; development and implementation of appropriate land management practices is the only means to make the urban growth sustainable (Zhao, 2010). The modern-day techniques like remote sensing play a vital role in assisting the decision makers to take informed measures. A number of techniques are available for mapping the built-up areas (Bhatti & Tripathi, 2014; Lo & Choi, 2004; Powell, Roberts, Dennison, & Hess, 2007), and one of the basic methods to study the urban sprawl is to examine the temporal variations in the land across heterogeneous geographical areas (Wilson, Clay, Martin, Stuckey, & Vedder-Risch, 2003; Zeng, Sui, & Li, 2005). The land managers also take help from the simulation models that assist is estimating the future urban growth (Zanganeh Shahraki et al., 2011). However, urban sprawl varies in different areas depending upon the land conversion patterns, which involves a number of factors. Zhao (2010) found that the socioeconomic factors and attitudes to residential plots influence the transportation patterns, which consequently affects the land development in peri-urban areas. Trip distances also influence the urban sprawl (Kenworthy & Laube, 1996). The significance of the socioeconomic and physical factors to observe the dynamics of urban growth have been established by other researchers as well (Longley & Mesev, 2000). Serra, Pons, and Saurí (2008) used biophysical and socioeconomic factors to identify the driving factors responsible for land use change, whereas Almeida *et al.* (2005) used infrastructure and socioeconomic indicators for analyzing the probability for urban growth. Population growth, employment and change in built-up area have also been used as the indicators of urban growth (Fulton, Pendall, Nguyen, & Harrison, 2001).

Most of the studies have employed methods to determine the LULC conversions considering the area under examination as a single region (Mundia & Murayama, 2010; Zanganeh Shahraki et al., 2011). For instance, Irwin and Bockstael (2004) studied land changes in the residential/urban areas only, whereas Martinuzzi, Gould, and Ramos González (2007) used the built-up density in both urban and rural areas to determine the urban sprawl. A few solely concentrated on transforming the rural landscapes/peri-urban to urban forms (Shi et al., 2012). However, Miller and Grebby (2014) focused on sprawl by classifying their study area into urban (densely built-up), peri-urban (houses having gardens) and rural (green and pervious surfaces) areas but did not consider the factors governing land changes. They confirmed that the peri-urban areas have more rapid growth than the urban or rural.

The land use types and driving factors exhibit a non-linear relationship in both space and time and encompass many factors that may be categorized as biophysical (topography, slope, geographic conditions, etc.), infrastructure (roads, business centers, industries, etc.) and socioeconomic (population growth, population density, employment opportunities, etc.). The artificial neural network (ANN) framework carefully handles such non-linear relationships (Thapa & Murayama, 2012) and the most commonly used is the multilayer perceptron neural network (MLPNN) (Hu & Weng, 2009; Kavzoglu & Mather, 2003). Based on the network developed from land classes and driving factors, the ANN efficiently determines the areas that are likely to change, however it could not decide how much to change. Thus, specific land demands can be computed through empirical or dynamic models such as Markov chain (MC), system dynamic, etc., which present quite reliable estimates (Luo, Yin, Chen, Xu, & Lu, 2010; Ti-yan, 2007). The accuracy of modeling also needs to be tested for validation purposes; area under the receiver operating characteristic curve (AUC) method has been efficient at evaluating the LULC change model's accuracy (Peterson, Papeş, & Soberón, 2008; Pontius & Schneider, 2001). This method compares the actual LULC change with the one computed through the model, and quantifies the level of agreement between the both.

Nevertheless, urban growth is an inevitable process which, if meticulously addressed, can be considered as an indicator of social and economic development in any region. This study presents an approach to handle the variable growth dynamics of urban and peri-urban areas (within a metropolitan region) in a simulation model. Four exploratory scenarios are developed based on two different LULC change rates and two land development conditions. The specific objectives of this study comprise: (1) exploratory analyses of model inputs and outputs at metropolitan scale and its subsets (urban and peri-urban); (2) multi-scale simulation of different exploratory scenarios and accuracy assessment using the actual land changes; and (3) future LULC prediction through the devised modeling approach to examine the spatio-temporal dynamics in the study area.

2. Study area

Lahore, the capital city of the province of Punjab, Pakistan, was selected as the study area for this research. The city, also termed as "city district Lahore", is stressed in terms of rapid urbanization with total population of around 9.16 million (2013 estimates) where 82% resides in the urban and the rest in peri-urban areas (Bureau of Statistics, 2013). The city is administratively divided into 10 towns (including a cantonment) and covers an area of around 179000 hectares (Figure 1A). The towns are further subdivided into 150 union councils (UCs), where 122 are urban and the rest are peri-urban/rural (Bureau of Statistics, 2013). Population in the city district has increased manyfold during the past decades, and clearly indicates the trend of urbanization (Figure 1B). The urban and peri-urban populations in 1951 in the city district were around 0.85 and 0.28 million, respectively (Population Census Organization, 1998); the difference in both was around 0.57 million at that time, which increased to 5.9 million in 2013 (Bureau of Statistics, 2013).

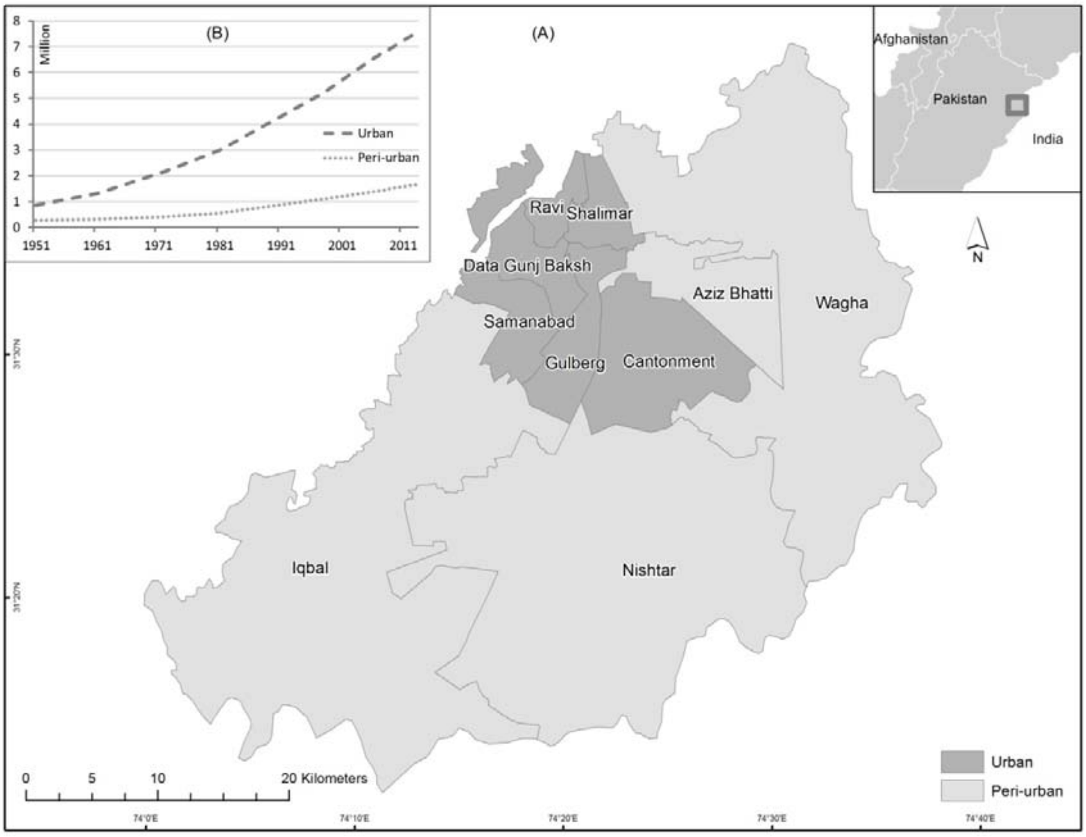

Figure 1. (A) Study area map showing the city district Lahore (urban and peri-urban towns), Pakistan and (B) population growth in the study area from 1951 to 2013 (Bureau of Statistics, 2013; Population Census Organization, 1998).

The study area was divided into two zones, urban and peri-urban. The towns having more than 70% of the area covered by the urban UCs were classified as urban, whereas the rest were categorized as peri-urban. With an area of around 26400 ha, the urban zone included Cantonment, Data Gunj Baksh, Gulberg, Ravi, Samanabad and Shalimar towns, while the peri-urban zone comprised Aziz Bhatti, Iqbal, Nishtar and Wagha towns covering around 152600 ha of area.

3. Methods of data

3.1. Datasets

A variety of datasets were used in this study for modeling the urban growth in city district Lahore at three spatial scales, city district, urban and peri-urban. Images acquired from landsat thematic mapper and operational land imager satellites (30 m spatial resolution) were processed through a hybrid classification approach (a mix of supervised and unsupervised classification) (Castellana, D'Addabbo, & Pasquariello, 2007; Lo & Choi, 2004) to obtain the LULC maps of 1999, 2011 and 2013. Based on the focus of this study (urban growth modeling) and dominant geographical features in the study area, five LULC classes were mapped, which included agriculture (cropland and area used for agricultural activities), bare (unutilized land and open spaces), built-up (residential, commercial, industrial and transportation), vegetation (trees, shrubs and grasslands) and water (open water features, streams, canals and river) (Anderson, 1976).

Other datasets mainly comprised the driving factors of LULC change, which were acquired from different sources including Advanced Spaceborne Thermal Emission and Reflection Radiometer, OpenStreetMap, The Urban Unit, Lahore and the reports on District Census of Lahore, Multiple Indicator Cluster Survey and Punjab Development Statistics. They were classified into three categories namely biophysical, infrastructure and socioeconomic to analyze their influence on urban growth (Table 1). Appropriate pre-processing was carried out to prepare each dataset for further analyses. The Euclidean distance method was used to obtain the distance to streams/canals, housing schemes, roads, city center, built-up, built-up change areas (1999-2011), railway lines, hospitals and schools (Batisani & Yarnal, 2009). All input datasets were prepared at 30 m spatial resolution to be consistent with that of the LULC maps.

3.2. Change analysis and selection of land transitions

The land change modeler (LCM) module of IDRISI Selva software was used to simulate the land changes; the LULC maps of 1999 and 2011 were used to prepare the model, whereas the map of 2013 was used for the validation of prediction results. Future projections were made for 2021 and 2035. Cross tabulation method was used to develop a transition matrix to show the change in the state of each LULC class over the period from 1999 to 2011 in the city district, urban and peri-urban areas, separately. The land transitions involving a considerable amount of land change area, and significant to this study (urban growth modeling), were then selected for modeling.

Table 1. Driving factors and notations.

Category	Notation	Driving Factor
Biophysical	Elev	Surface elevation
	Slop	Surface slope
	DStr	Distance to streams/canals
Infrastructure	DHSc	Distance to housing schemes
	DRd	Distance to roads
	DCC	Distance to city center
	DBU	Distance to built-up
	DBUC	Distance to areas changed to built-up during 1999-2011
	DR	Distance to railway lines
	DH	Distance to hospitals
	DS	Distance to schools
	ImpDW	Percentage households having access to improved drinking water
	DWPre	Percentage households having drinking water access on premises
	ImpSa	Percentage households having access to improved sanitation facilities
	DisWW	Percentage households having facilities for proper disposal of wastewater
	DisSW	Percentage households having facilities for proper disposal of solid waste
Socioeconomic	PopG	Annual population growth rate
	PopD	Population density
	Lit	Literacy rate
	Emp	Percentage population employed
	OH	Percentage population having ownership of the house

3.3. Exploratory analysis and selection of driving factors

A critical aspect of urban growth modeling is the selection of driving factors that can be associated to the LULC change (Thapa & Murayama, 2010). Cramer's *V* analysis, which quantitatively measures the association between two variables, was performed for the selection of driving factors. This analysis was used to test whether or not a driving factor explained a particular land transition. The Cramer's *V* value is computed by Equation 1.

$$V = \sqrt{\frac{\emptyset^2}{\min(k-1, r-1)}} \tag{1}$$

Where $\emptyset$ is the mean square contingency coefficient, k is the number of columns and r is the number of rows (Acock & Stavig, 1979). The value of Cramer's *V* ranges from 0 to 1, where a higher value indicates greater association between the land class and driving factor being tested and vice versa. A value of *V* greater than or equal to 0.15 implies the usefulness of that particular driving factor, while a value above 0.4 suggests a good association (Eastman, 2012). For each selected land transition, the factors returning Cramer's *V* value greater than or equal to 0.15 were selected. In line with the objectives of this study, the driving factor sets were prepared at three spatial extents, city district, urban and peri-urban, for all selected transitions.

3.4. Sensitivity analysis and development of sub-models

After the selection of land transitions and pertinent driving factors, a sensitivity analysis was conducted using the relative operating characteristic (ROC) method through logistic regression to finalize the land transitions appropriate for development of sub-models (Mozumder & Tripathi, 2014; Oñate-Valdivieso & Bosque Sendra, 2010). The ROC value ranges from 0 to 1, where values higher than 0.5 indicate some association between the maps of reality and suitability, and values close to 1 indicate a strong fit between the two maps (Eastman, 2012). Only the transitions having ROC value greater than or equal to 0.75 were considered appropriate, and were selected for the development of sub-models. These transitions were grouped into sub-models, where a sub-model shares a common set of driving factors (Eastman, 2012; Geneletti, 2013; Mozumder & Tripathi, 2014). Subsequently, four sub-models, each were determined for the city district, urban and peri-urban areas to generate the transition potential maps.

The contribution of biophysical, infrastructure and socioeconomic driving factors to the land transitions at the city district, urban and peri-urban scales was also analyzed. For each land transition, the contribution percentage (CP) of the three types of driving factors was computed separately by Equation 2.

$$CP = \frac{DF_S}{DF_T} \times 100 \qquad (2)$$

Where DF_S is the number of driving factors selected from a particular domain (biophysical, infrastructure or socioeconomic) for any transition, and DF_T is the total number of driving factors in that particular domain. The value of CP was used to analyze the association of each driving factor domain with the land transitions.

3.5. Transition potential modeling and determination of transition rates

A separate transition potential map was generated for each land transition modeled through the MLPNN in LCM. For each sub-model consisting "X" number of transitions, "2X" example classes were fed, half of which comprised the transition samples and the rest were the persistent samples. A network of neurons was created between the "2X" example classes and the corresponding driving factors, where the web of connections comprised the sets of weights that were initially determined randomly by the MLPNN (Eastman, 2012; Mozumder & Tripathi, 2014). These weights were adjusted during each iteration to obtain an accurate set to generate a multivariate function. Out of the total samples selected by the MLPNN, 50% were used for training, whereas the rest were used for the validation of neural network. Modeling accuracies were tested for the sub-models and separate transition potential maps were generated for all land transitions at the city district, urban and peri-urban scales.

Two transition rates were considered in this study, the first one (R1) was derived by MC prediction (Bell, 1974; Geneletti, 2013), which considers that the type and rate of future land transitions will be the same as in the past. The second rate (R2) was determined considering a more rapid rise in urbanization, and increasing the Markovian rate by 50% for the land transitions to built-up (Geneletti, 2013).

3.6. Transition scenarios and LULC prediction

Two growth scenarios were considered in this study; the first one (S1) was the business as usual in which no constraints or preferences were set on the future land transitions, whereas the second one (S2) was based on the restrictions and preferences for future LULC in the study area. The map of constraints/incentives was prepared for S2 that comprised four classes: prohibited (educational institutes, transportation areas like airports, parks and recreational areas, areas around streams and railway lines, and floodplain), disfavored (waterlogged, vegetated and water areas), neutral (all areas except prohibited, disfavored or favored) and favored (preset and planned housing schemes). Integer values of 0, 0.5, 1 and 2 were assigned to these classes respectively. During the change prediction process, this map is multiplied by the transition potential maps where the numeric value of each class in the constraints/incentives map restricts, decreases or increases the transition potential in the respective areas (Eastman, 2012).

The MLPNN transition potential maps and the transition rates (R1 and R2) were used to produce the prediction maps of 2013 for both scenarios (S1 and S2). All four possible exploratory scenarios, R1S1, R1S2, R2S1 and R2S2, were considered to produce the prediction maps at three scales, city district, urban and peri-urban (total 12 maps).

3.7. Model validation and prediction of 2021 and 2035

All prediction maps of 2013 were compared with the actual LULC map of 2013 to examine their accuracy; AUC method was used which determines how well a continuous surface predicts the locations given in the distribution of actual LULC change (Eastman, 2012). The best approach (modeling at the city district or urban/peri-urban scale) was determined based on the AUC values, and the prediction maps of 2021 and 2035 were generated for all four exploratory scenarios using the selected approach. The selection of these years for prediction was based on fact that the local development authorities have prepared a master plan of 2021 (Jamal, Mazhar, & Kaukab, 2012; Nadeem, Haydar, Sarwar, & Ali, 2013; NESPAK-LDA, 2004), and are in process of preparing one of 2035 (Dawn, 2012; LDA, 2012; Nadeem et al., 2013) for the city district Lahore. The results of this study could be useful for the concerned departments and may help improving the future planning.

4. Results and discussion

4.1. LULC change analysis

The urban growth dynamics in the study area were examined at city district, urban and peri-urban scales for the period from 1999 to 2011; Figure 2 and Table 2 shows the LULC changes and the cross tabulation results respectively.

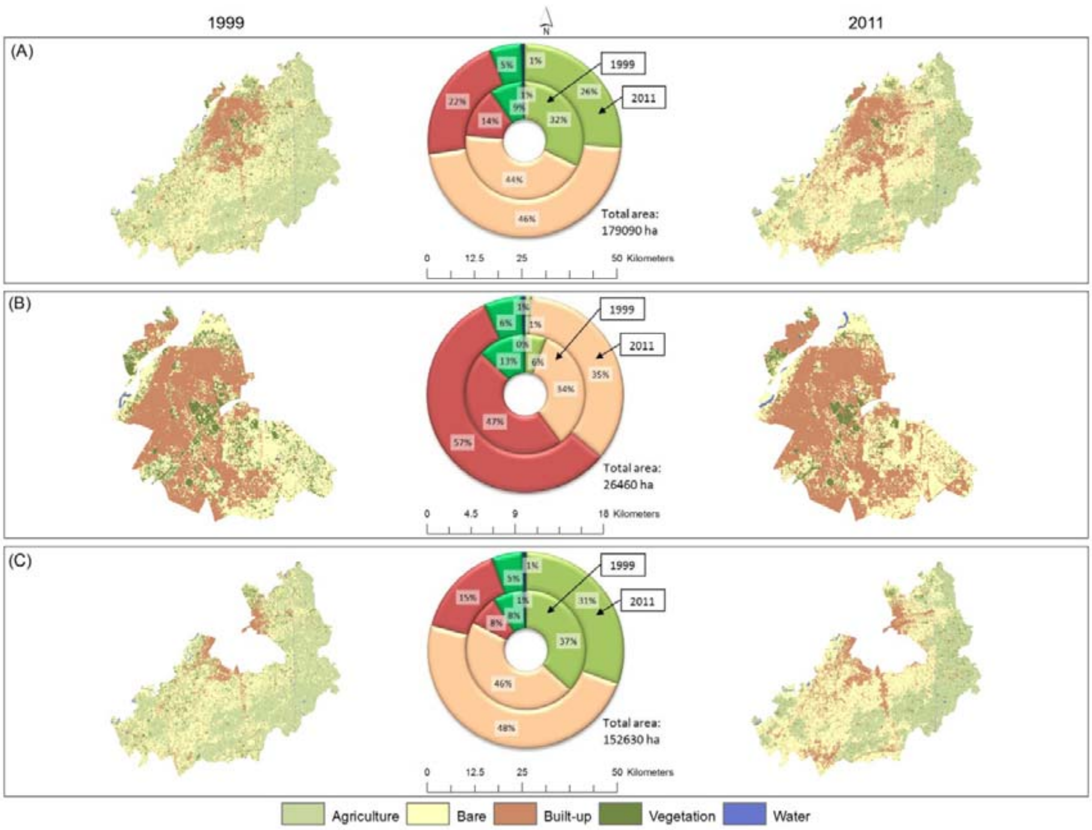

Figure 2. LULC maps and area graphs of 1999 and 2011 at the (A) city district, (B) urban and (C) peri-urban scales.

Table 2. Cross tabulation of LULC changes between 1999 and 2011 (in hectares).

	From (1999)				
To (2011)	Agriculture	Bare	Built-up	Vegetation	Water
City District					
Agriculture	43,239.78	1,875.42	1.98	1,533.96	180.99
Bare	**8,289.36**	66,521.43	6.30	**7,324.29**	399.15
Built-up	**4,103.82**	**6,404.40**	25,104.87	**3,858.39**	141.30
Vegetation	1,917.09	**3,750.03**	2.97	3,486.06	69.57
Water	73.53	441.00	0.63	65.61	293.58
Urban					
Agriculture	102.24	35.55	0.45	128.34	2.88
Bare	**659.16**	7,628.67	1.35	**878.22**	57.24
Built-up	**566.64**	**1,120.32**	12,448.71	**1,044.99**	7.56
Vegetation	215.91	124.83	2.07	1,298.79	2.43
Water	6.66	84.06	0.00	14.94	26.28
Peri-urban					
Agriculture	43,137.54	1,839.87	1.53	1,405.62	178.11
Bare	**7,630.20**	58,892.76	4.95	**6,446.07**	341.91
Built-up	**3,537.18**	**5,284.08**	12,656.16	**2,813.40**	133.74
Vegetation	1,701.18	**3,625.20**	0.90	2,187.27	67.14
Water	66.87	356.94	0.63	50.67	267.30

Numbers in bold indicate significant changes, and their corresponding transitions are considered in this study.

At city district scale, it was observed that the majority of built-up area existed towards the north and northwestern parts in 1999, which extended towards the south by 2011 basically due to the rise in population (Figure 2A). The built-up area almost doubled during this period where major contributors of land were bare, agriculture and vegetation (Table 2). Examining the study area at urban scale, the majority of built-up area expansion was found in the northwestern and eastern parts where it increased from 47% of the total urban area in 1999 to 57% of that in 2011 (Figure 2B). The major contributors to the built-up area, along with bare, were vegetation and agriculture (Table 2). A significant reduction in the agricultural and vegetated areas was also observed. At the peri-urban scale, a significant rise in built-up area (around 84%) was observed towards the northern (adjoining the urban areas) and southern parts during 1999 and 2011 (Figure 2C). The major land contributors to built-up area were bare, agriculture and vegetation (Table 2). Vegetation and agricultural area reduced by around 39% and 17%, respectively, during this period in the peri-urban zone.

These results indicate that the significant land transitions were different at different spatial scales within the same metropolitan region. Thus, separate land transitions were selected for urban growth modeling at the city district, urban and peri-urban scales which involved four LULC classes and included: (1) agriculture to bare, agriculture to built-up, bare to built-up, bare to vegetation, vegetation to bare and vegetation to built-up in city district; (2) agriculture to bare, agriculture to built-up, bare to built-up, vegetation to bare and vegetation to built-up in urban; and (3) agriculture to bare, agriculture to built-up, bare to built-up, bare to vegetation, vegetation to bare and vegetation to built-up in peri-urban areas.

4.2. Analysis of driving factors and land transitions

The association between the four significant land classes and twenty-one driving factors was checked quantitatively through Cramer's *V* values (Table 3). Instead of considering the overall *V*, the values of individual LULC classes were examined as they provide a better indication of the association between driving factors and land classes (Eastman, 2012).

Table 3. Cramer's *V* coefficients showing the quantified association between selected LULC classes and the driving factors (city district, urban and peri-urban scales).

Driving factors	Overall *V*			Agriculture			Bare			Built-up			Vegetation		
	City district	Urban	Peri-urban	City district	Urban	Peri-urban	City district	Urban	Peri-urban	City district	Urban	Peri-urban	City district	Urban	Peri-urban
Biophysical															
Elev	0.0167	0.036	0.1214	0.0132	0.0305	0.0442	0.0038	0.003	0.003	0.1132	0.1115	**0.1628**	0.0045	0.0148	0.0046
Slop	0.0167	0.036	0.1214	0.0132	0.0305	0.0442	0.0038	0.003	0.003	0.1132	0.1115	**0.1628**	0.0045	0.0148	0.0046
DStr	0.1201	0.1753	0.1424	**0.2057**	0.1214	**0.2364**	0.1037	**0.2256**	0.1358	0.1289	**0.2013**	**0.1599**	0.0497	**0.1651**	0.0682
Infrastructure															
DHSc	0.2115	0.1493	0.2264	**0.3733**	0.1291	**0.3819**	**0.2102**	**0.1672**	**0.2364**	**0.2147**	0.115	**0.2641**	0.1132	**0.2028**	0.1115
DRd	0.2627	0.2009	0.2269	**0.4058**	**0.1586**	**0.3617**	**0.1821**	**0.2489**	**0.2076**	**0.3919**	**0.3216**	**0.2918**	0.0886	0.1029	0.1071
DCC	0.0843	0.0211	0.0427	0.0968	0.024	0.0363	0.0546	0.04	0.0328	**0.1556**	0.093	0.0769	0.0291	0.027	0.0277
DBU	0.2888	0.1911	0.2743	**0.5063**	0.0379	**0.4833**	**0.2014**	**0.2577**	**0.2542**	**0.3902**	**0.324**	**0.3462**	0.0577	0.1135	0.0636
DBUC	0.1884	0.1382	0.1971	**0.356**	0.0286	**0.3589**	**0.1664**	0.0982	**0.1946**	**0.2011**	0.138	**0.2273**	0.0326	0.0897	0.053
DR	0.2156	0.1495	0.1893	**0.3516**	0.0617	**0.3168**	**0.1832**	**0.2553**	**0.2132**	**0.2853**	**0.2264**	**0.1994**	0.0509	0.126	0.0672
DH	0.2827	0.2165	0.2187	**0.4044**	0.0951	**0.3441**	**0.229**	**0.3761**	**0.2069**	**0.4508**	**0.3913**	**0.29**	0.0903	0.0753	0.1115
DS	0.1832	0.251	0.1438	**0.2162**	0.095	**0.1811**	0.0818	**0.3537**	0.0623	**0.3063**	**0.4042**	**0.1977**	0.0227	0.0911	0.0203
ImpDW	0.0313	0.1704	0.0394	0.0132	0.0305	0.0442	0.0266	**0.3388**	0.0213	0.0566	**0.2714**	0.0534	0.0245	0.0854	0.0467
DWPre	0.0534	0.1705	0.0709	0.0693	0.0246	0.0993	0.0165	**0.339**	0.0745	0.0794	**0.2713**	0.0705	0.0513	0.0899	0.08
ImpSa	0.1673	0.0175	0.1467	**0.3098**	0.0061	**0.2695**	0.1412	0.0332	**0.1831**	**0.1862**	0.0331	0.1265	0.0542	0.0006	0.0662
DisWW	0.0167	0.1214	0.036	0.0324	0.0305	0.0242	0.0711	0.003	0.0543	0.1208	**0.153**	0.0469	0.0067	0.0148	0.0123
DisSW	0.1078	0.073	0.0794	**0.1897**	0.13	0.1444	0.0451	0.053	0.074	**0.1512**	0.092	0.1185	0.0136	0.015	0.02
Socioeconomic															
PopG	0.225	0.1538	0.1621	**0.2736**	0.012	**0.1987**	**0.1519**	**0.3054**	0.0252	**0.4229**	**0.245**	**0.3013**	0.0094	0.0849	0.02
PopD	0.2698	0.1913	0.1938	**0.353**	0.0563	**0.2585**	**0.193**	**0.3169**	0.1419	**0.4753**	**0.3815**	**0.3272**	0.0399	0.1199	0.0521
Lit	0.0837	0.1214	0.0519	0.1375	0.1015	0.0912	0.0017	0.003	0.0247	0.1282	**0.157**	0.0636	0.0229	0.0145	0.0214
Emp	0.0581	0.0871	0.0475	0.1042	0.0326	0.0838	0.0773	**0.1688**	0.0817	0.0377	0.1147	0.0103	0.0324	0.0733	0.0234
OH	0.125	0.0801	0.0783	**0.2052**	0.0588	0.143	0.0208	0.1319	0.0858	**0.1967**	0.146	0.0706	0.0122	0.0015	0.0245

The notations used for driving factors are explained in Table 1. Numbers in bold indicate acceptable relationship (*V* value of 0.15 or higher).

The majority of driving factors had an acceptable association (Cramer's *V* value > 0.15) with agriculture at the city district and peri-urban scales; however, only distance to roads was found to have a better association with agriculture at the urban scale. This implies that the changes in agricultural areas in the urban zone are mainly related to the physical accessibility factor (transportation through roads). Different sets of driving factors were found to be associated with the bare and built-up areas at the three spatial scales. A decent relationship was observed between distance to housing schemes and built-up change areas with built-up class at the city district and peri-urban scales, however, this association was weak in the urban areas. Similar kinds of differences were also found in several other driving factors related to the bare class in urban areas, implicating that the factors governing land changes are different from each other in the urban and peri-urban zones.

An interesting finding was the weak association of driving factors with vegetation at all three spatial scales. This could be attributed mainly to the irregular changes in vegetation that might have resulted due to variability in weather conditions or some other pertinent factors. Since the focus of this study was to examine changes in the built-up areas and not the vegetation, the factors selected could not completely explain the changes in vegetated areas. Nevertheless, an important thing deduced was that the driving factors governing land changes are different at different spatial scales, hence implying the need to model and simulate the land transitions separately for urban and peri-urban areas.

A sensitivity analysis was conducted by the ROC method through logistic regression for final selection of land transitions for modeling (Table 4). The majority of selected land transitions had a strong relationship with the associated driving factors at the city district, urban and peri-urban scales (high ROC value). However, three transitions returned ROC value less than 0.75 (weak relationship with the driving factors), which were therefore dropped (Table 4).

Table 4. ROC values showing the level of association between selected land transitions and the driving factors (city district, urban and peri-urban scales).

City district			Urban			Peri-urban		
Sub-model	LULC transition	ROC	Sub-model	LULC transition	ROC	Sub-model	LULC transition	ROC
1	Agriculture to bare	0.8771	-	Agriculture to bare	**0.6957**	1	Agriculture to bare	0.8681
2	Agriculture to built-up	0.9989	1	Agriculture to built-up	0.9737	2	Agriculture to built-up	0.9985
3	Bare to built-up	0.9922	2	Bare to built-up	0.9629	2	Bare to built-up	0.9928
3	Vegetation to built-up	0.9862	3	Vegetation to bare	0.7845	3	Vegetation to built-up	0.9913
4	Bare to vegetation	0.846	4	Vegetation to built-up	0.946	4	Bare to vegetation	0.8687
-	Vegetation to bare	**0.6587**				-	Vegetation to bare	**0.59**

Numbers in bold indicate weak relationship (ROC value less than 0.75), and their corresponding transitions are discarded.

The percentage association of each category of driving factors (biophysical, infrastructure and socioeconomic) with the selected land transitions was analyzed separately at city district, urban and peri-urban scales. The land transitions at city district scale were explained mainly by the infrastructure related driving factors, followed by the socioeconomic and biophysical categories (Figure 3A), indicating that the majority of land transitions were taking place as a result of changes in infrastructure and socioeconomic conditions. The trend was slightly different in the urban zone where all selected land transitions, except for bare to built-up, were explained chiefly by the infrastructure related factors, followed by the socioeconomic and biophysical ones (Figure 3B). The bare to built-up transition was explained majorly by the socioeconomic related driving factors implying that this transition was more sensitive to the changes in selected socioeconomic variables. In peri-urban areas, the majority of land changes were found to be related to the variability in biophysical and infrastructure related factors (Figure 3C).

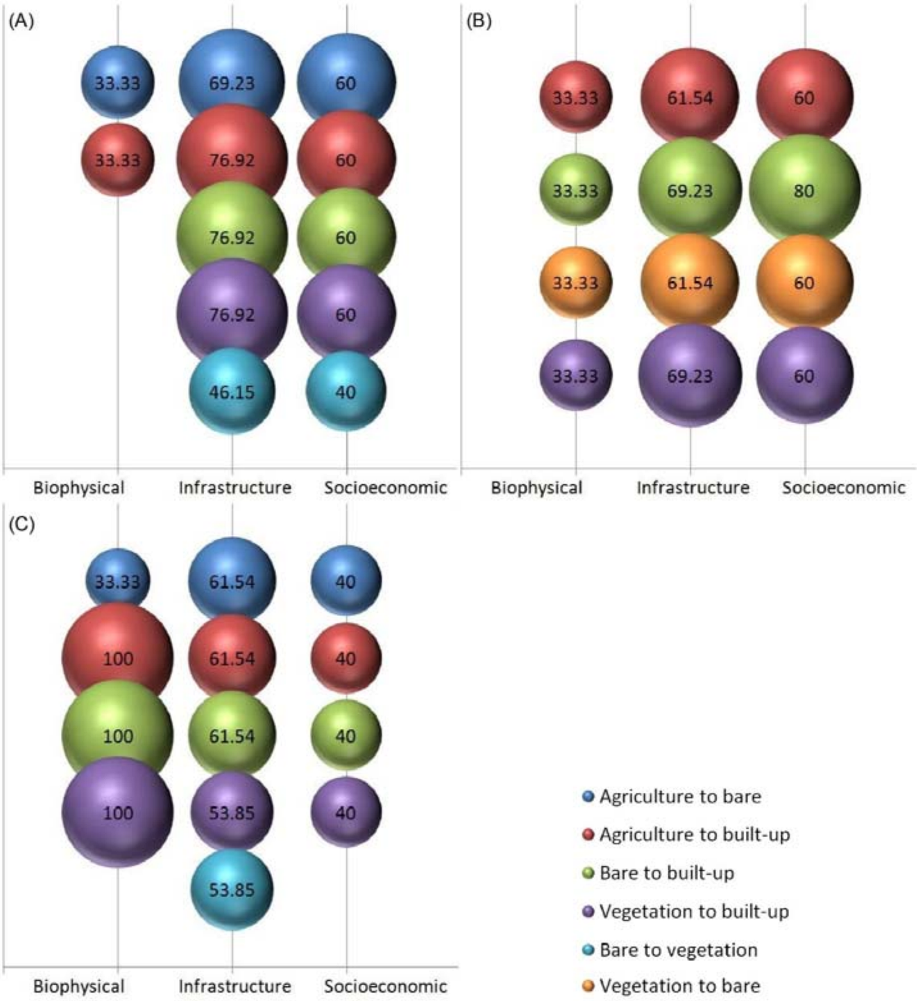

Figure 3. Percentage of driving factors in each category contributing to the land transitions in the (A) city district, (B) urban and (C) peri-urban areas. The size of the circle represents the percentage, bigger means higher and vice versa.

The land transitions were grouped into four sub-models, each for the city district, urban and peri-urban areas, where a sub-model comprised a common set of driving factors (Table 4). At city district scale, the accuracies of about 71, 74, 67 and 72 percent were obtained from transition potential maps of the sub-models 1, 2, 3 and 4, respectively. The accuracies of outputs from the urban area sub-models 1, 2, 3 and 4 were around 79, 73, 81 and 75 percent, respectively, whereas for peri-urban areas, the accuracies of around 84, 72, 79 and 87 percent were obtained from the sub-models 1, 2, 3 and 4, respectively. Examining the averages of sub-model accuracies at the city district (71%), urban (77%) and peri-urban (81%) scales, it could be inferred that the land transitions were explained better by the driving factors at urban and peri-urban scales than the ones at the city district scale. This implies that the transition potential maps at urban and peri-urban scales were more suitable for modeling land changes than the city district ones.

4.3. Prediction results and model validation

The prediction maps of 2013 were generated at city district, urban and peri-urban scales for the four exploratory scenarios (R1S1, R1S2, R2S1 and R2S2), and prediction accuracies were evaluated through the AUC method. Figures 4A-C shows the prediction maps and Figure 4D shows the actual LULC map of 2013. Figure 5 shows the area statistics comparison between all predicted maps and the actual LULC map of 2013.

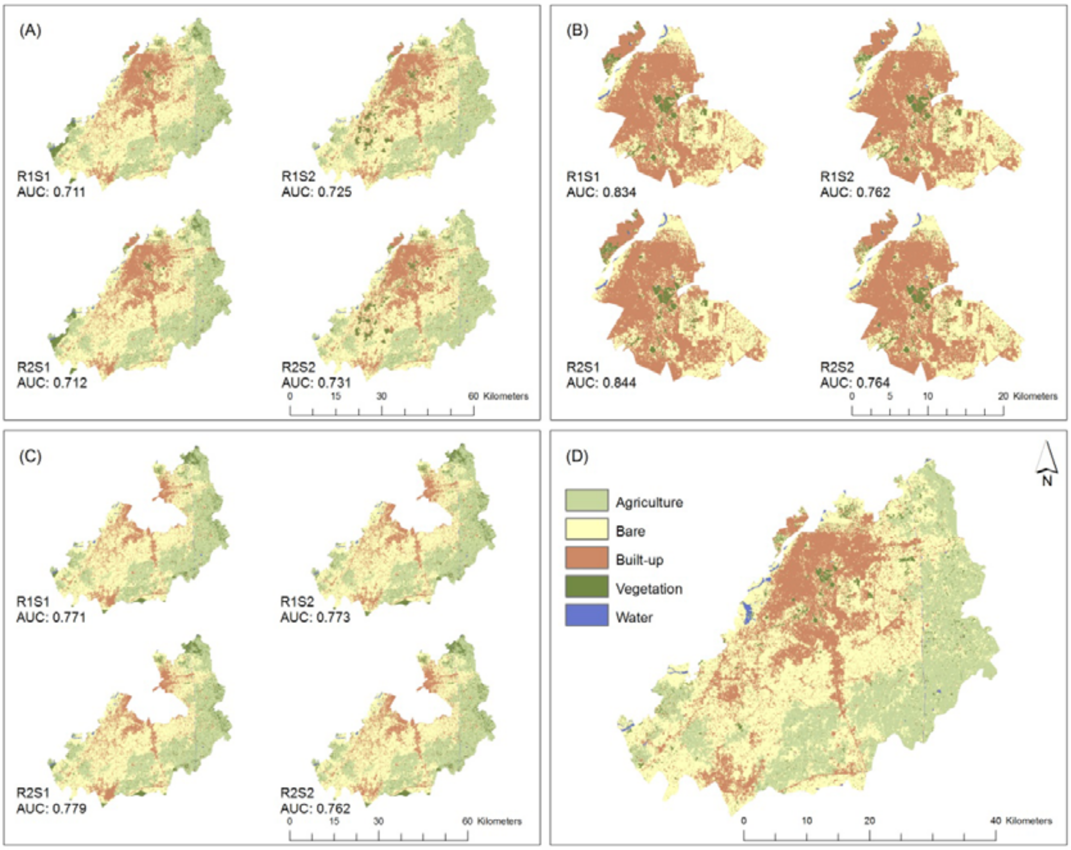

Figure 4. Prediction maps of 2013 with the AUC values for all exploratory scenarios at the (A) city district, (B) urban and (C) peri-urban scales, and (D) the actual LULC map (2013).

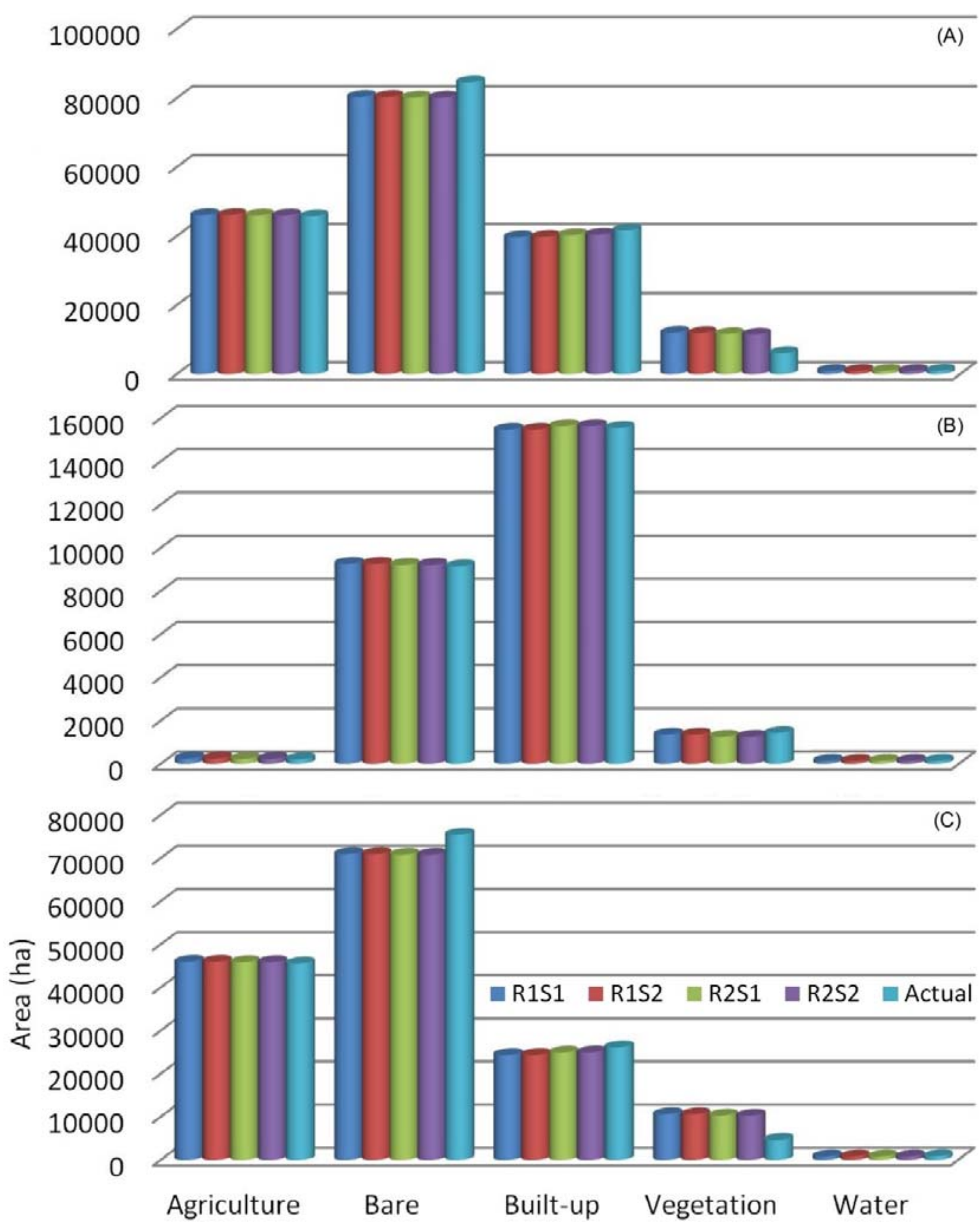

Figure 5. Predicted and actual land areas of 2013 for all scenarios at the (A) city district, (B) urban and (C) peri-urban scales.

At city district scale, the prediction maps of all the scenarios indicated an expansion of built-up area mainly towards the southern parts of the metropolitan region (Figure 4A). The areas of all LULC classes were almost similar in R1S1 and R1S2, and were also comparable in R2S1 and R2S2 (Figure 5A). The areas of agriculture and water classes were predicted with a decent accuracy, when compared to the actual ones in 2013, whereas that of vegetation was on the higher side in the predictions. This deviation can be attributed to the weak association between driving factors and the vegetation class that was observed during the analysis of driving factors. The areas of bare and built-up classes were slightly less in all predicted maps compared to the actual ones. However, the accuracies obtained in terms of AUC values of 0.711, 0.725, 0.712 and 0.731 for the exploratory scenarios R1S1, R1S2, R2S1 and R2S2, respectively, were quite decent. At urban scale, an increase in the built-up area was observed towards the city center and the east (Figure 4B). The area statistics reveal that the predictions for all scenarios were quite close to the actual land areas (Figure 5B). In addition, the AUC values of 0.834, 0.762, 0.844 and 0.764 for the exploratory scenarios R1S1, R1S2, R2S1 and R2S2, respectively, also indicated that the prediction maps were quite accurate in comparison with the actual LULC map of 2013. An expansion in the built-up area was observed mainly towards the south when examined at the peri-urban scale (Figure 4C). The area statistics indicated slight differences in the predicted areas of bare, vegetation and built-up classes compared to the actual ones in 2013 (Figure 5C). However, the AUC values of 0.771, 0.773, 0.779 and 0.762, respectively, for the R1S1, R1S2, R2S1 and R2S2 scenarios implied that these differences were not significant and the predictions were reasonable.

Comparing the AUC values of prediction maps at different scales, it can be deduced that the predictions made at the urban and peri-urban scales were more accurate than the city district ones. This finding is in line with the results of sensitivity analysis (ROC values) and the MLPNN model accuracy statistics. The R2S2, R2S1 and R2S1 scenarios returned the highest accuracies at city district, urban and peri-urban scales, respectively. Since the prediction results are more accurate with R2S1 scenario at urban and peri-urban scales, it can be implied that the land conversions, at present, are taking place at a high rate without considering the restrictions or preferences for land transitions, thus signifying the need for appropriate land management.

4.4. Prediction of 2021 and 2035

The prediction maps of 2021 and 2035 were generated using the subset approach by modeling the urban and peri-urban areas separately for the R1S1, R1S2, R2S1 and R2S2 scenarios. However, for each scenario, the outputs from both the subsets were combined to present the whole metropolitan region (city district Lahore) (Figure 6).

The built-up area is expected to expand mainly towards the south and east of the metropolitan region by 2021 (Figure 6A), and is likely to further extend towards the west by 2035 (Figure 6B). Majority of the land transition towards the south and east is expected to occur at the cost of agricultural area. The area statistics in Figure 6 show the estimates of the increase in built-up and decrease in agricultural areas during the period from 2021 to 2035.

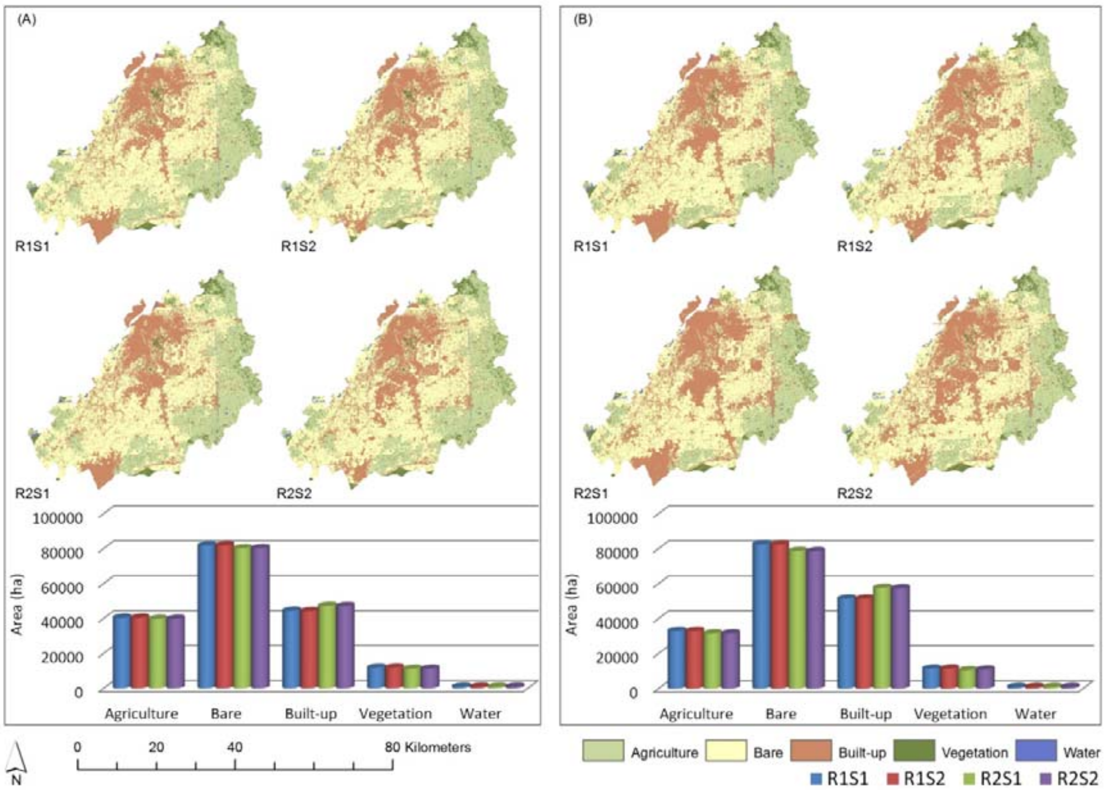

Figure 6. Prediction maps and area graphs of (A) 2021 and (B) 2035 for all exploratory scenarios.

Considering the R2S1 scenario to prevail in the future, as identified through the 2013 prediction (Section 4.3), the expansion and densification of built-up area is expected to be quite high around the urban areas, towards the west, south and east of the study area. Figure 7 shows the predicted LULC areas in 2021 and 2035 for R2S1; the statistics suggest a reduction in the agriculture and bare areas whereas the built-up area is expected to rise significantly.

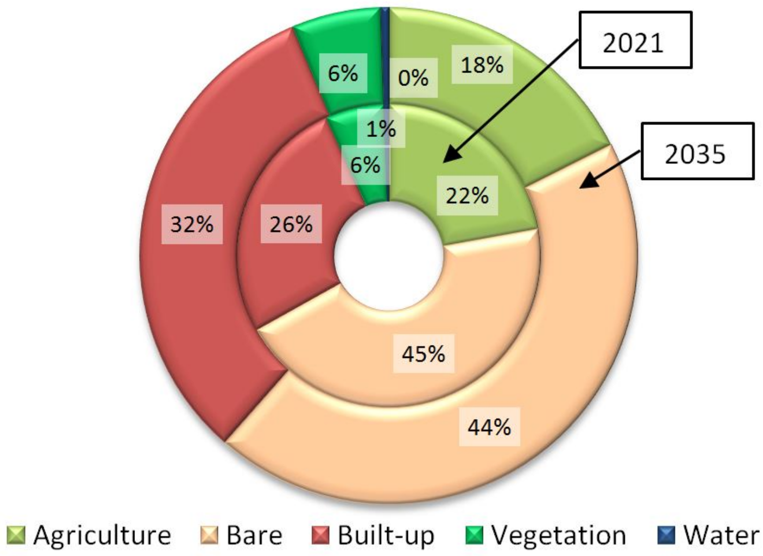

Figure 7. Predicted LULC areas of 2021 and 2035 for R2S1 scenario.

5. Conclusions

The LULC maps of 1999 and 2011 were examined at different scales, and it was found that the major land transitions varied in the urban and peri-urban zones within the metropolitan region. Moreover, the factors governing land changes were dissimilar for the same land transitions in both the zones. This finding was in conformity with the results discussed by Thapa & Murayama (2010), and suggested to consider multiple scales for analysis and modeling. The MLPNN modeling accuracies and the AUC values of the prediction maps of 2013 derived at multiple scales (city district, urban and peri-urban) verified this inference. These findings signify the need to develop careful understanding of the factors of land change in different zones within a metropolitan region; the planners need to develop separate land management strategies for urban and peri-urban areas.

The use of more than one growth scenarios for investigating the LULC changes has been demonstrated in several research studies (Geneletti, 2013; Mozumder & Tripathi, 2014; Oñate-Valdivieso & Bosque Sendra, 2010). The study of multiple growth scenarios contributed to this particular research in two ways; (1) they helped in understanding the present growth dynamics in the study area through comparison of the results of different growth scenarios with the actual LULC, and (2) assisted in examining the impacts of implementing different growth scenarios, especially the ones related to restricting/promoting LULC changes, on the future LULC conditions. The results imply that the proposed approach, by considering the differences in growth dynamics of the urban and peri-urban areas and integrating the various growth scenarios, could be useful to model and predict the urban growth in a metropolitan region.

Acknowledgements

The authors gratefully acknowledge the support from the Asian Institute of Technology, Thailand, and the Japanese Government for carrying out this research. We would also like to thank the Earth Resources Observation and Science Center, United States Geological Survey for providing Landsat satellite data free of charge for this study, and The Urban Unit, Lahore, Bureau of Statistics, Lahore, Department of City and Regional Planning (CRP), University of Engineering & Technology (UET), Lahore and Metropolitan Wing, Lahore Development Authority (LDA), Lahore for their support and providing the secondary data for this research. The authors would also like to thank the reviewers for their insightful comments and valuable suggestions.

References

Acock, A. C., & Stavig, G. R. (1979). A Measure of Association for Nonparametric Statistics. *Social Forces*, *57*(4), 1381–1386. http://doi.org/10.1093/sf/57.4.1381

Almeida, C. M. De, Monteiro, A. M. V., Câmara, G., Soares Filho, B. S., Cerqueira, G. C., Pennachin, C. L., & Batty, M. (2005). GIS and remote sensing as tools for the simulation of urban land-use change. *International Journal of Remote Sensing*, *26*(4), 759–774. http://doi.org/10.1080/01431160512331316865

Anderson, J. R. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data, Issue 964 (Google eBook)*. U.S. Government Printing Office. Retrieved from http://books.google.com/books?hl=en&lr=&id=dE-ToP4UpSIC&pgis=1

Batisani, N., & Yarnal, B. (2009). Urban expansion in Centre County, Pennsylvania: Spatial dynamics and landscape transformations. *Applied Geography*, *29*(2), 235–249. http://doi.org/10.1016/j.apgeog.2008.08.007

Bell, E. J. (1974). Markov analysis of land use change—an application of stochastic processes to remotely sensed data. *Socio-Economic Planning Sciences*, *8*(6), 311–316. http://doi.org/10.1016/0038-0121(74)90034-2

Bhatti, S. S., & Tripathi, N. K. (2014). Built-up area extraction using Landsat 8 OLI imagery. *GIScience & Remote Sensing*, *51*(4), 445–467. http://doi.org/10.1080/15481603.2014.939539

Bureau of Statistics. (2013). *Punjab Development Statistics 2013*. Lahore.

Castellana, L., D'Addabbo, A., & Pasquariello, G. (2007). A composed supervised/unsupervised approach to improve change detection from remote sensing. *Pattern Recognition Letters*, *28*(4), 405–413. http://doi.org/10.1016/j.patrec.2006.08.010

Dawn. (2012, July 24). LDA to prepare development plan-2035. Lahore. Retrieved from http://www.dawn.com/news/736735/lda-to-prepare-development-plan-2035

Deng, J. S., Wang, K., Hong, Y., & Qi, J. G. (2009). Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. *Landscape and Urban Planning*, *92*(3-4), 187–198. http://doi.org/10.1016/j.landurbplan.2009.05.001

Eastman, J. (2012). IDRISI Selva Tutorial. *Idrisi Production, Clark Labs-Clark University*. Retrieved from http://scholar.google.com/scholar?hl=en&q=IDRISI+Selva.+Clark-Labs,+Clark+University&btnG=&as_sdt=1,5&as_sdtp=#0

Fulton, W., Pendall, R., Nguyen, M., & Harrison, A. (2001). *Who sprawls most?: How growth patterns differ across the US*. Retrieved from http://www.brookings.edu/~/media/research/files/reports/2001/7/metropolitanpolicy fulton/fultoncasestudies.pdf

Geneletti, D. (2013). Assessing the impact of alternative land-use zoning policies on future ecosystem services. *Environmental Impact Assessment Review*, *40*, 25–35. http://doi.org/10.1016/j.eiar.2012.12.003

Hu, X., & Weng, Q. (2009). Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. *Remote Sensing of Environment*, *113*(10), 2089–2102. http://doi.org/10.1016/j.rse.2009.05.014

Irwin, E. G., & Bockstael, N. E. (2004). Land use externalities, open space preservation, and urban sprawl. *Regional Science and Urban Economics*, *34*(6), 705–725. http://doi.org/10.1016/j.regsciurbeco.2004.03.002

Jaeger, J. a. G., Bertiller, R., Schwick, C., & Kienast, F. (2010). Suitability criteria for measures of urban sprawl. *Ecological Indicators*, *10*(2), 397–406. http://doi.org/10.1016/j.ecolind.2009.07.007

Jamal, T., Mazhar, F., & Kaukab, I. S. (2012). Spatio-temporal residential growth of Lahore city. *Pakistan Journal of Science*, *64*(4), 293–295. Retrieved from http://www.paas.com.pk/images/volume/pdf/1023735823-faiza mazhar.pdf

Kavzoglu, T., & Mather, P. M. (2003). The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing*, *24*(23), 4907–4938. http://doi.org/10.1080/0143116031000114851

Kenworthy, J. R., & Laube, F. B. (1996). Automobile dependence in cities: An international comparison of urban transport and land use patterns with implications for sustainability. *Environmental Impact Assessment Review*, *16*(4-6), 279–308. http://doi.org/10.1016/S0195-9255(96)00023-6

LDA. (2012). *Terms of reference for preparation of integrated strategic development plan for Lahore region 2035*. Lahore. Retrieved from http://www.urbanunit.gov.pk/ISDP/TORs LAHORE LDA_ISDP35_July12-2012.pdf

Lo, C. P., & Choi, J. (2004). A hybrid approach to urban land use/cover mapping using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) images. *International Journal of Remote Sensing*, *25*(14), 2687–2700. http://doi.org/10.1080/01431160310001618428

Longley, P. A., & Mesev, V. (2000). On the measurement and generalisation of urban form. *Environment and Planning A*, *32*(3), 473–488. http://doi.org/10.1068/a3224

Luo, G., Yin, C., Chen, X., Xu, W., & Lu, L. (2010). Combining system dynamic model and CLUE-S model to improve land use scenario analyses at regional scale: A case study of Sangong watershed in Xinjiang, China. *Ecological Complexity*, *7*(2), 198–207. http://doi.org/10.1016/j.ecocom.2010.02.001

Martinuzzi, S., Gould, W., & Ramos González, O. M. (2007). Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. *Landscape and Urban Planning*, *79*(3-4), 288–297. http://doi.org/10.1016/j.landurbplan.2006.02.014

McGee, T. G. (1995). Metrofitting the Emerging Mega-Urban Regions of ASEAN: An Overview. In T. G. McGee & I. M. Robinson (Eds.), *The Mega-urban Regions of Southeast Asia* (p. 384). UBC Press. Retrieved from http://books.google.com/books?id=BOHsEzpJhx4C&pgis=1

Miller, J. D., & Grebby, S. (2014). Mapping long-term temporal change in imperviousness using topographic maps. *International Journal of Applied Earth Observation and Geoinformation*, *30*, 9–20. http://doi.org/10.1016/j.jag.2014.01.002

Mozumder, C., & Tripathi, N. K. (2014). Geospatial scenario based modelling of urban and agricultural intrusions in Ramsar wetland Deepor Beel in Northeast India using a multi-layer perceptron neural network. *International Journal of Applied Earth Observation and Geoinformation*, *32*, 92–104. http://doi.org/10.1016/j.jag.2014.03.002

Mundia, C. N., & Murayama, Y. (2010). Modeling Spatial Processes of Urban Growth in African Cities: A Case Study of Nairobi City. *Urban Geography*, *31*(2), 259–272. http://doi.org/10.2747/0272-3638.31.2.259

Nadeem, O., Haydar, S., Sarwar, S., & Ali, M. (2013). Consideration of environmental impacts in the integrated maste plan for Lahore-2021. *Pakistan Journal of Science*, *65*(3), 310–317. Retrieved from http://www.paas.com.pk/images/volume/pdf/2129395290-25-13 - Final Copy for Publication.pdf

NESPAK-LDA. (2004). *Integrated Master Plan for Lahore-2021*. Lahore.

Oñate-Valdivieso, F., & Bosque Sendra, J. (2010). Application of GIS and remote sensing techniques in generation of land use scenarios for hydrological modeling. *Journal of Hydrology*, *395*(3), 256–263. http://doi.org/10.1016/j.jhydrol.2010.10.033

Peterson, A. T., Papeş, M., & Soberón, J. (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, *213*(1), 63–72. http://doi.org/10.1016/j.ecolmodel.2007.11.008

Pontius, R. J., & Schneider, L. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment*, *85*, 239–248. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167880901001876

Population Census Organization. (1998). *Population Census Data of Pakistan*. Islamabad.

Powell, R., Roberts, D. A., Dennison, P. E., & Hess, L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment*, *106*(2), 253–267. http://doi.org/10.1016/j.rse.2006.09.005

Serra, P., Pons, X., & Sauri, D. (2008). Land-cover and land-use change in a Mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. *Applied Geography*, *28*(3), 189–209. http://doi.org/10.1016/j.apgeog.2008.02.001

Shi, Y., Sun, X., Zhu, X., Li, Y., & Mei, L. (2012). Characterizing growth types and analyzing growth density distribution in response to urban growth patterns in peri-urban areas of Lianyungang City. *Landscape and Urban Planning*, *105*(4), 425–433. http://doi.org/10.1016/j.landurbplan.2012.01.017

Thapa, R. B., & Murayama, Y. (2010). Drivers of urban growth in the Kathmandu valley, Nepal: Examining the efficacy of the analytic hierarchy process. *Applied Geography*, *30*(1), 70–83. http://doi.org/10.1016/j.apgeog.2009.10.002

Thapa, R. B., & Murayama, Y. (2012). Scenario based urban growth allocation in Kathmandu Valley, Nepal. *Landscape and Urban Planning*, *105*(1-2), 140–148. http://doi.org/10.1016/j.landurbplan.2011.12.007

Ti-yan, S. (2007). Study on Spatio-Temporal System Dynamic Models of Urban Growth. *Systems Engineering-Theory & Practice*, *27*(1), 10–17. Retrieved from http://scholar.google.com/scholar?q=Study+on+spatio-temporal+system+dynamic+models+of+urban+growth+ti-yan&btnG=&hl=en&as_sdt=0,5#2

Wilson, J. S., Clay, M., Martin, E., Stuckey, D., & Vedder-Risch, K. (2003). Evaluating environmental influences of zoning in urban ecosystems with remote sensing. *Remote Sensing of Environment*, *86*(3), 303–321. http://doi.org/10.1016/S0034-4257(03)00084-1

Zanganeh Shahraki, S., Sauri, D., Serra, P., Modugno, S., Seifolddini, F., & Pourahmad, A. (2011). Urban sprawl pattern and land-use change detection in Yazd, Iran. *Habitat International*, *35*(4), 521–528. http://doi.org/10.1016/j.habitatint.2011.02.004

Zeng, H., Sui, D. Z., & Li, S. (2005). Linking Urban Field Theory with Gis and Remote Sensing to Detect Signatures of Rapid Urbanization on the Landscape: Toward a New Approach for Characterizing Urban Sprawl. *Urban Geography*, *26*(5), 410–434. http://doi.org/10.2747/0272-3638.26.5.410

Zhao, P. (2010). Sustainable urban expansion and transportation in a growing megacity: Consequences of urban sprawl for mobility on the urban fringe of Beijing. *Habitat International*, *34*(2), 236–243. http://doi.org/10.1016/j.habitatint.2009.09.008